

Multidimensional assessment of R&D performance: evidence from the 2014 pilot evaluation exercise of Russian public research organizations¹

Konstantin Fursov*, Gleb Kuzmin**

*kfursov@hse.ru

Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics, 20 Myasnitskaya Street, Moscow, 101000 (Russia)

**gkuzmin@hse.ru

Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics, 20 Myasnitskaya Street, Moscow, 101000 (Russia)

Research evaluation recently became a widely disseminated exercise aimed in the end of the day at improving the cost efficiency of public funding of national R&D sectors. In November 2013, the Government of the Russian Federation initiated a national evaluation exercise of public research organizations (PROs) to provide information basis for development of S&T policies aimed at increasing effectiveness and strengthening the role of R&D performing institutions in economic and social development. The aim of this paper is that of providing an approach for multidimensional assessment of R&D performance based on quantitative data derived from the national evaluation exercise, specifically looking at its applicability and limitations for further analysis and preliminary differentiation of PROs as well as for use in policymaking.

Introduction

Research evaluation of public institutes is considered important for measuring the performance of R&D sector, first of all, in terms of improving the cost efficiency of public funding. Various studies (e.g. see Luwel et al., 1999; Senker 2001; Coccia, 2004) show a growing interest in evaluating national scientific performance with the use of different metrics and models.

In November 2013, the Government of the Russian Federation initiated a national evaluation exercise of public research organizations (PROs) to provide information basis for development of S&T policies aimed at increasing effectiveness and strengthening the role of R&D performing institutions in economic and social development. This exercise is to be the second one in the post-soviet academic history. The first initiative of this kind was taken in 2009 – 2011, however its results did not lead to any substantial changes in policy instruments or institutional organisation of the national R&D sector. One of the major shortcomings was a narrow agency-level character of the evaluation procedure with the leading role of administrative authorities interested in preserving research organizations under their jurisdiction from any kind of critical assessment, intervention or reformation and therefore using different performance indicators. Despite the limitations of the first round of evaluation exercise, it stimulated Agora processes (Barré, 2001, 2005), i.e. rise of public debates on the use of S&T indicators for research assessment and overall innovation development. In order to rise the efficiency of the national R&D performing organisations, the Russian Government initiated a new program of research evaluation. The distinctive features of this initiative as

¹ This work was supported by the Ministry of Education and Science of the Russian Federation in 2014-2015 (Grant No 16.602.21.0004).

assigned in the corresponding regulation documents envisage an open and interagency approach for research evaluation based on combination of quantitative and qualitative (peer-review based) assessment procedures. Moreover, organisations entering the evaluation process should be compared only within their reference groups that provide the benchmarks in terms of research field and basic research functions. An evaluation is conducted once in 5 years while data is collected yearly, so reference group configuration and preliminary assessment is possible in a shorter period.

The aim of this paper is that of providing an approach for multidimensional assessment of R&D performance based on quantitative data, specifically looking at its applicability and limitations for further analysis and preliminary differentiation of organizations as well as for use in policymaking.

Sources and methods

Data for this study is derived from the PRO Federal Monitoring database maintained and updated by the Ministry of Education and Science of the Russian Federation (<http://www.sciencemon.ru/>) in November 2014. Reference year is 2013. The initial dataset included information on different characteristics of R&D performance for the 1033 PROs submitted their reports for monitoring and further evaluation. Additionally arithmetic and logical controls were conducted for key variables to construct a final sample covering 626 observations. Research outputs considered for further assessment were journal articles, conference proceedings and book chapters indexed in Web of Science and Scopus; books and scientific periodicals, designs, registered intellectual property rights (IPRs) and others. Additionally a number of financial results such as income from technology transfer, S&T services and contractual works provided were taken into account.

Preliminary correlation analysis demonstrated strong significant cohesion between basic research and publication output. At the same time, as seen from the table 1, organizations implementing applied research and experimental development are likely to be the main providers of S&T services and have weak linkages to IPR protection. In other words organizations with identical balance between different types of research activities may have different results and vice versa. Therefore, they cannot be included into a single and homogenous reference group and compared. Similarly, it could be shown that other characteristics such as type of institution, legal status of an organization or its size are weak to differentiate R&D outputs.

Table 1 – Pearson correlation matrix

Variable		Publications		IPRs		S&T services	
		abs*	norm*	abs	norm	abs	norm
R&D expenditure		-					
Basic research	value	0.81**		0.07**		0.13**	
	% of total		0.27**		-0.03**		-0.15**
Applied research	value	0.06**		0.27**		0.62**	
	% of total		-0.23**		0.04**		0.09**
Experimental development	value	0.03**		0.24**		0.70**	
	% of total		-0.13**		0.00**		0.13**

* abs = absolute values, and norm = normalised values.

** (Sig. = 0.99)

In order to construct the reference groups for PROs a two-step algorithm was implemented. At the first step, research organizations were divided into groups by field of science (the first level of the OECD FOS classification (OECD, 2002) in line with the specific research areas

mentioned by the organizations. As a result, an organization may refer to more than one field of science. At the second step, within each field of science, a research profile of an organization describing its orientation towards one or several social functions, i.e. knowledge generation, technology development or provision of S&T services, and consequently corresponding with certain types of research outputs was identified. Herewith a research profile was determined with the use of the following key indicators:

- A. The number of books or scientific periodicals and papers indexed in Web of Science, Scopus or other specialized database (maximum value from one of the databases representative for an organization) per 100 researchers.
- B. Number of IPRs registered in the Russian Federation or abroad as well as the number of issued design documentation per 100 researchers.
- C. Income from contractual R&D, S&T services provided and technology transfer per total R&D personnel of an organization.

As seen from the above, to reduce scale effects each indicator (A – C) reflecting certain type of outputs is divided by the number of employees, mainly involved in its acquisition. While researchers tend to be mostly engaged in publications and IPR production, financial results require involvement of different categories of R&D personnel.

Consequently, an organization refers to a particular research profile reflecting its bent for certain research function and corresponding output, if one or more of the indicators mentioned above (A – C) is not zero and equal or exceeds the median value for the field of science. It allows identification up to eight research profiles (see figure 1). An intersection of field of science and a research profile constitutes a reference group.

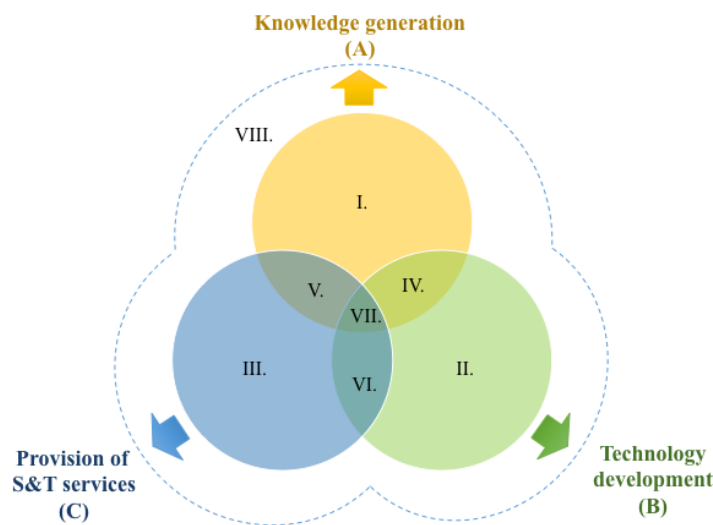


Figure 1 – Research profiles of organisations

For further differentiation of organisations within a reference group, a second median value for the same set of core performance indicators of a research profile in the relevant field of science was taken. For instance, an indicator A will be core for the research profiles I, IV, V and VII. Then, for research profiles IV and V indicators B and C will also be core. The research profile VIII should be specifically mentioned. It includes organizations demonstrating lower level of research productivity by all indicators. At the same time, it may have some other characteristics, not included to the initial selection criteria.

According to the requirements mentioned in the related legal documents regulating overall evaluation procedure (see: <http://www.sciencemon.ru/legal/acts/postanovlenie-pravitelstva-rf-ot-1-noyabrya-2013-979/>), the leadership within the reference group is distributed as follows:

- 1st category (leading institutions) = median + 25%
- 3rd category (loosing research functions) = median – 25%
- 2nd category (stable research organisations) = others

General distribution of organisation by reference groups and performance categories is provided below as a possible solution obtained from the available sample and implementation of the above-mentioned criterion. To distinguish between 2nd and 3rd performance categories within the VIII research profile it was suggested to use additional criterion measured as a share of private funding in the intramural R&D expenditure. The role of private financing was discussed in (Coccia, 2004). In other words if an organization demonstrates poor performance level, but not "at public expense" (share of private funding is equal or above the median level within the relevant field of science) it falls into the 2nd category. Otherwise, it is associated with institutions loosing their research functions and requires further peer-review based expertise.

Preliminary results

Achieved sample allowed identification of eight research profiles in six fields of science, which provides adequate representation of reference groups and thereby – the statistical significance of the derived threshold values. In the present structure natural sciences, medical and agricultural sciences as well as technology studies are well represented. Somewhat lower concentration of organizations is seen in in the social sciences and humanities (table 2).

Each research profile is characterized by a unique configuration of resources and results that can not apply the same evaluation criteria and thresholds to the whole population of organizations. In particular, the resulting distribution naturally include confluent profiles of technology developers and service providers (VI) in the field of humanities. Similar misbalances can be seen in other fields.

Table 2 – Model distribution of PROs by reference groups and performance categories

Field of Science	Category	Total	Research profile							
			I.	II.	III.	IV.	V.	VI.	VII.	VIII.
1. Natural Sciences	Total	342	34	40	48	48	39	33	51	49
	1-st category		14	14	17	3	5	6	1	X
	2-nd category		15	13	14	41	34	25	50	28
	3-rd category		5	13	17	4	0	2	0	21
2. Engineering and Technology	Total	186	22	21	27	25	19	20	27	25
	1-st category		10	8	8	2	1	1	1	X
	2-nd category		5	4	11	19	18	18	26	14
	3-rd category		7	9	8	4	0	1	0	11
3. Medical and Health Sciences	Total	142	15	19	20	20	19	15	17	17
	1-st category		5	8	6	1	3	2	0	X
	2-nd category		6	6	10	18	16	11	17	6
	3-rd category		4	5	4	1	0	2	0	11
4. Agricultural Sciences	Total	138	9	26	15	12	23	19	12	22
	1-st category		4	8	7	3	3	1	2	X
	2-nd category		1	8	7	9	18	18	10	12
	3-rd category		4	10	1	0	2	0	0	10
5. Social Sciences	Total	76	9	7	10	8	10	7	11	14
	1-st category		3	2	2	1	1	1	0	X
	2-nd category		4	2	8	7	9	6	11	3
	3-rd category		2	3	0	0	0	0	0	11
6. Humanities	Total	32	5	3	9	4	5	0	2	4
	1-st category		2	1	0	1	1	0	0	X
	2-nd category		1	1	9	2	4	0	2	0
	3-rd category		2	1	0	1	0	0	0	4

Different fields of science demonstrate different functional roles. In particular, in the field of medical sciences most of the organizations appear in the group of technology developers (II and IV profiles, mostly first and second performance categories), agricultural sciences also dominated by the developers as well (specifically in the first category). Social sciences are represented mostly in the VIII profile that is characterized by a high number of ulterior (unexpressed) results mostly due to the high number of libraries and infrastructure elements historically holding status of research organizations. The latter is also true for the technical and technological and natural science areas having a number of pilot plants, experimental stations, etc. The second reason for over-representation of the VIII profile is multiple research orientation by fields of science and inability within the existing framework to divide measurable types of outputs by corresponding fields of science, which is one of the limitations of the proposed approach.

Discussion and further research

The proposed approach for reference group identification and assessment could be easily replicated in different samples. However, it has several significant limitations.

The results of the described distribution are largely determined by the boundary values set for the each reference group. Changing these values may lead to significant changes in the whole picture. In the proposed model the thresholds are set automatically, depending on the behavior of all scientific organizations in a sample. Accordingly, any change in the sample can make a difference in the assessment for an individual organization. Such changes may be both productive (e.g. increased performance indicators of some organizations over time can worsen the situation for others) or counterproductive (for instance, through the introduction of unusual objects in the sample for which the proposed indicators would not be relevant).

Secondly, the method does not exclude the possibility for an organization to change its category due to minimal change in one of the performance indicators. In other words, when two organizations in the same field of science demonstrate about the same level of performance and one of them lies on the threshold separating research profiles, then they may be included in different categories that is logically incorrect, taking into account their almost complete identity. Thus, the organizations holding leadership positions in the same pure profiles (I to III) may lose their advantage in relation to one of the mixed profiles (IV to VII), i.e. improving value of an additional indicator, it may lose its leading position by another one.

Third, the current approach does not allow further differentiation of R&D outputs by field of science. Therefore, performance is measured by the overall output of an organisation that may significantly distort the assessment results in case of its multidisciplinary specialization.

In the end, it should be mentioned that, such an exercise can be taken only as a preliminary and rough assessment the results of which though may be used for further detailed analysis and peer-review bases evaluation. Further research is aimed at overcoming the described constraints and searching for a better-optimized solution.

References

Barré, R., 2001. The Agora model of innovation systems: S&T indicators for a democratic knowledge society. *Research Evaluation* 10, 13–18.

Barré, R., 2005. S&T indicators for policy making in a changing science-society relationship, in: *Handbook of Quantitative Science and Technology Research*. Springer, 115–131.

Coccia, M., 2004. New models for measuring the R&D performance and identifying the productivity of public research institutes. *R&D Management* 34, 267–280.

Luwel, M., Noyons, C.M. and Moed, F., 1999. Bibliometric assessment of research performance in Flanders: policy background and implications. *R&D Management*, 29, 2, 33–141.

OECD. 2002. *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development*. OECD, 67.

Senker, J., 2001. Changing organisation of public sector research in Europe – implications for benchmarking human resources in RTD. Paper prepared for Human resources in RTD session of the contribution of European socio-economic research to the benchmarking of RTD policies in Europe. Conference, Brussels, March 15–16.